DOCUMENT RESUME

ED 199 687 CS 005 982

AUTHOR Stratton, Beverly D.; And Others
TITLE Use of Case History Data for the Development of Equations in Predicting High Risk, Reading Disabled Students.
PUB DATE Apr 81
NOTE 25p.; Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Birth Order; Correlation; *High Risk Students; Multiple Regression Analysis; Predictor Variables; *Reading Diagnosis; *Reading Difficulties; *Reading Research; *Siblings

ABSTRACT

Demographic data on 92 subjects identified as having reading problems were used to develop equations useful in identifying high risk, reading disabled students. Multiple linear regression analysis of the data indicated that reading disability (1) had a significant positive relationship with birth order and number of siblings; (2) had a positive but not significant relationship with chronological age and evidence of learning problems in the family; and (3) had a negative but not significant relationship with IQ. IQ was found to have negative but not significant relationships with birth order and number of siblings. The interpretation of the data suggested a possible link between reading disability and a lack of environmental stimulation. The importance of early identification and appropriate remediation was also apparent. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED199687

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

USE OF CASE HISTORY DATA FOR THE DEVELOPMENT OF EQUATIONS IN PREDICTING HIGH RISK READING DISABLED STUDENTS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Beverly D. Stratton
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Beverly D. Stratton
The University of Akron

Joan Charlton-Seifert
The University of Akron

Isadore Newman
The University of Akron

Mary Maher-Boehnlein
Cleveland State University

A Paper Presented at the Annual Meeting
of the
American Educational Research Association
Los Angeles, California
April 17, 1981

CS005902

Introduction

The child's proficiency in academic learning depends upon a number of related variables. Those identified by specialists include the proper functioning of the sense organs, central nervous system, skeletal-muscular system, and biochemical system. Other variables which have also been identified include the child's social competence, emotional stability, cultural background, and educational experiences.

Approximately eighty-five percent of all school children have been able to succeed in the developmental reading programs currently in use in the schools. However, about fifteen percent of the students have not been able to master fundamental reading skills and, so, have become educationally disabled in other subjects which are primarily dependent upon reading ability (Kaluger and Kolson, 1969). Nearly every classroom contains some students whose reading levels fall below their potential (Gulliland, 1974).

Various writers emphasize the primacy of one variable in accounting for student's learning/reading problems. Bond and Tinker (1973) stated that educational experiences are the most common cause of academic difficulties, while Delacato (1966) claimed that central nervous dysfunction is the most frequent etiology. Biochemical etiology is stressed by Watson (1972), while Jencks (1972) cited cultural background as the predominate variable affecting learning. Bettelheim (1973), on the other hand, concluded that the key casual factor behind both general and specific learning problems is emotional disturbance.

Because the home produces the first, most insistent impact on a child, its importance has long been recognized. Studies by Altus (1965), Cicirelli (1967), Hansen (1969), Hodges and Balow (1961), Oberlander and Jenkins (1967), and Otto (1965) investigated the significance of the family constellation in terms of birth order and its relationship to academic and reading success. The evidence has not been consistent, however.

Bakwin (1973); Critchley (1970); Gillingham and Stillman (1960); Hermann (1959); Michael-Smith, Morgenstern, and Karp (1970); Silver (1971); and Symmes and Rapoport (1972) believed that reading disability is genetically linked and therefore inherited. They considered the presence of similar difficulties, i.e., skills in working with printed and written symbols, in other family members to be an important factor for consideration in diagnosis.

The Purpose of the Study

The purpose of this study was threefold: 1) to identify potential demographic variables available in case histories which identify high risk, reading disabled students, 2) to determine whether these variables were important predictors of a high risk, reading disabled population, and 3) to demonstrate how case history data can be coded in order to make data collected in a clinical setting readily accessible for further study and research.

Subjects and Design

Data was collected on ninety-two children who had been identified by teachers and/or parents as having a reading problem.

Each child had been administered an individual intelligence test, an informal oral reading test, and a standardized silent reading test. Determination of individual student's degree of reading disability was based upon that child's mental potential for learning to read. Reading expectancy levels were estimated through utilization of the Harris Reading Expectancy Chart (1975).

Since there were no active variables in this research study, the study, by definition, was expost facto (Newman, Deitchman, Burkholder, Sanders, and Ervin, 1976). It is not legitimate technically, therefore, to infer causation when the design is expost facto (Newman, et. al., 1976). The statistical questions being asked in this study were related to whether there was a significant relationship between certain independent variables and a criterion. Multiple Linear Regression was employed to analyze the data.

A set of 12 hypotheses were tested relating to the relationship between seven independent variables (sex = x_2, x_3; I.Q. = x_4; chronological age = x_5; birth order = x_6; number of siblings in family excluding student = x_7; evidence of learning problems in other family members = x_8; marital status of the parents = x_9 to x_{15}) and their ability to predict reading disability (x_1).

Statistical Analysis

The statistical procedure employed in analyzing the hypotheses was Multiple Linear Regression. This procedure was selected for the following reasons: 1) Multiple Linear Regression is the general case of the least sum of square solution (Draper and Smith, 1966); Kerlinger, 1973; McNeil, Kelly and McNeil, 1975; Newman, 1976; Ward and Jennings, 1973). Other statistical techniques such as chi square, t and F tests are calculated on the basis of the least sum of square solution. When a researcher could use any of the special cases of the least squares solution, the more general case would also be appropriate (Newman, 1976). 2) Regression is more flexible in that it allows one to write linear models which specifically reflect the research questions. 3) With regression a researcher can test the relationship between categorical variables, between categorical and continuous variables, or between continuous variables. There is no need to categorize variables that are continuous in nature as required by traditional analysis of variance; therefore, one would not lose degrees of freedom and power (Kerlinger, 1973; McNeil, Kelly, and McNeil, 1975; Newman, 1976).

All of the hypotheses were tested using the F test. The formula for the F test is as follows:

$$F\ (df_n,\ df_d) = \frac{(R^2f - R^2r)/df_n}{(1-R^2f)df_d}$$

Where:

R^2f = the proportion of observed criterion variance accounted for by the full linear model.

R^2r = the proportion of observed criterion variance accounted for by the restricted linear model.

df_n = the difference between the number of linearly independent vectors in the full model and the number of linearly independent vectors in the restricted model.

df_d = the number (N) of observations minus the number of linearly independent vectors in the full model (McNeil, Kelly, McNeil, 1975, p. 102-103).

The full Model (R^2f) reflects the research hypothesis and contains more information. Testing the two models reflects statistical hypothesis.

Two-tailed tests of significance were used to test the relationships of those variables where the direction was uncertain. One-tailed tests of significance were used where the direction of the correlation was certain based upon previous research and experience.

The use of .05 level of significance was a subjective decision made by the researcher. The consequences of making a Type I error, rejecting a true null hypothesis, was not so serious as to warrant a more stringent confidence level.

The power of the study was calculated for medium effect size differences from the following formula (Newman and Benz, 1979):

$$L = f^2 v$$

where: f^2 = effect size (.15)

$v = df_2 (N-m_1)$

$u = df_1 (m_1-m_2)$

With an N of 92 and an alpha level of .05 the estimate of power was equal or greater than .60. That is, at least 60 times out of 100 a medium effect difference could be detected if it existed.

Probability statements for the following relationships were generated from the correlation matrix:

$r_{x_1, x_4} \neq 0$

$r_{x_1, x_5} \neq 0$

$r_{x_1, x_6} \neq 0$

$r_{x_1, x_7} \neq 0$

$r_{x_1, x_8} \neq 0$

$r_{x_4, x_6} \neq 0$

$r_{x_4, x_7} \neq 0$

There was a need for the researchers to keep the alpha level constant across the seven comparisons since the comparisons were not independent of each other. The method used was $\frac{\text{alpha}}{N-1}$ (Newman and Fry, 1972). The procedure was to divide the

alpha level (.05) by the number of comparisons made minus one (7-1 = 6). The result, .00833, was the probability the F ratio had to meet in order to be significant at the .05 level for a one-tailed test. It should be noted that this corrected alpha (.00833) is an accurate but highly conservative, stringent estimate which tends to overcorrect (Newman, 1976).

Limitations of the Study

Following is a list of limitations that apply to this study. These must be kept in mind when interpreting the results of this study.

1. The independent variables could not be manipulated by the researchers to insure internal validity, causation, therefore, cannot be inferred.

2. The limitation of specific population is imposed by using existing data from children identified by teachers and/or parents as having a reading problem. The generalizability of the findings will be limited to students of the same type as the sample used in this study.

3. Past research demonstrated that there is no method which can precisely measure the degree of reading disability. The reading disability calculated for each individual in this study then are estimates.

Results and Conclusions

Table 1 summarizes the statistical findings of multiple linear regression as suggested by Newman (1973). The table includes the following information:

1. a statement of the research hypotheses;
2. an identification of the full and restricted models used to test the hypothesis;
3. the squared multiple correlation coefficient (R^2);
4. the degrees of freedom (df);
5. the alpha level needed for significance (Alpha);
6. the F-ratio (F);
7. the probability (p);
8. significance (S = significance - probability $<$ .05
   NS = nonsignificance - probability $>$ .05).

- -

Insert Table 1 here

- -

Since one of the major goals of research is to isolate the unique effects of variables, other competing explainers of the data must be eliminated (McNeil, Kelly, McNeil, 1975). Statistical control, through analysis of covariance, was the procedure used in the first six hypotheses to isolate the effects of certain variables on reading disability. In hypothesis 5, knowledge of birth order, number of siblings, and evidence of learning problems in the family accounted for approximately 9 percent of the criterion variance that is independent of what can be accounted for by the covariates, i.e., sex, IQ, chronological age, and marital status of the parents. This was statistically significant ($p = .0401$).

The correlational data suggested the following:

1. The relationship between reading disability and IQ ($r = -.0519$) is negative and not significant.

2. The relationship between reading disability and chronological age ($r = .1550$) is positive but not significant.

3. There is a significant positive relationship between reading disability and birth order ($r = .2720$).

4. There is a significant positive relationship between reading disability and number of siblings in family ($r = .2720$).

5. The relationship between reading disability and evidence of learning problems in the family ($r = .0198$) is positive but not significant.

6. The relationship between IQ and birth order ($r = -.1541$) is negative but not significant.

7. The relationship between IQ and number of siblings ($r = -.2394$) is negative but not significant.

Results and Conclusions

On the basis of the statistical analysis of the data provided by this study, the following conclusions are drawn:

1. Evidence of learning problems in the family appeared unrelated to the other variables individually;

2. Males with reading problems had significantly higher I.Q.'s than females ($r = .4129$);

3. Older children with reading problems had lower I.Q. scores ($r = -.3989$), thus pointing out the importance for early identification of high risk, reading disabled students;

4. Students living in a foster home had significantly lower I.Q.'s than other students in the study (r = .2335); and

5. First born children tended to have higher, but not significantly higher I.Q.'s.

Cross validation procedures were conducted in order to determine the stability of the regression weights and the R^2 values and also to develop the most parsimonious equation to determine the best and most stable predictors from sample to sample.

Cross validation procedures indicated that the regression weights were not stable. This was due to the high colinearity between predictor variables. This does not mean, however, that this set of predictors or variables will not produce a relatively stable R (a shrunken R) figure. However, these separate predictor variables should not be assigned apriori fixed weights since the weighting of these variables are likely to vary from sample to sample. In other words, the variables birth order and number of siblings in the family highly correlate (r = .91). In some samples birth order may take an extra weight. A suggestion would be to drop one of these variables from the predictor set. On the other hand, the variables of sex and IQ correlate moderately high (r = .41). One may not want to drop either of these variables from the predictor set because of the medium high r between these variables. The regression weights, however, are likely to fluctuate from sample to sample. It should be evident from these two examples that the higher the

intercorrelations between predictor variables the larger the variability of the regression weights (therefore the less stability from sample to sample). The important point here is that even though the weights may not be stable, they will be significantly predictive of Reading Disability. One would not expect the R^2 to be less than .26 in the population using Wherry's shrinkage estimate which is intended to measure the R^2 from sample to population (Uhl and Eisenberg, 1970).

Post Hoc Analysis

After visual inspection of the correlation matrix and consideration of the logic of the variables as a predictor set, it was decided to eliminate four variables, i.e., number of siblings and three variables dealing with the marital status of the parents and rerun the data. These variables were eliminated because of high intercorrelations with other predictor variables and/or they did not make sufficient predictive sense. Testing the full model against chance produced an R^2 of .14 ($p \leq .000009$) and an estimated shrinkage of .05. The estimate of power for medium effect size differences was calculated for the more parsimonious predictor set. With an N of 92 and an alpha level of .05. The estimate of power was equal to or greater than .67. That is, at least 67 times out of 100 a medium effect difference could be detected if it existed. It was felt by the researchers that this final predictive equation makes more empirical sense and it is this predictive equation which should be investigated further.

Educational Importance of the Study

Much information is available in case histories that can be used in a variety of ways. For example, this study identified demographic variables of importance in the development of equations to identify high risk, reading disabled students.

Reading diagnosticians need to be aware that when dealing with a clinic population, I.Q. scores alone are not significantly predictive of reading disability. In the population in the study, males demonstrated a higher I.Q. than females. This phenomena is not true of the general population. It may be that I.Q. scores of students in this study were more negatively affected due to reading problems. I.Q. tests have a tendency to measure certain reading abilities and many language abilities which may be attained through reading (Burmeister, 1974). Students who continue to experience reading problems may experience depressed I.Q. scores because they are not reading and this effect becomes increasingly strong over time if the reading difficulty is not ameliorated. This emphasizes the importance of early identification and appropriate remediation (Clarke, 1973).

A possible link between a lack of environmental stimulation as well as potential hereditary factors can be drawn from the relationship between birth order and reading disability. Birth order, which is related to the number of siblings in the family, was significantly related to reading disability.

All of the demographic variables in this study were identified from case histories. The ability of these variables to predict reading disability produced an r of approximately .44 which is highly significant ($p < .001$). This appears to indicate that these variables can identify and screen potentially high risk, reading disabled students.

REFERENCES

Altus, W. B. Birth order and scholastic aptitude. Journal of Consulting Psychology, 1965, 29, 202-205.

Bakwin, H. Reading disability in twins. Developmental Medicine and Child Neurology, 1973, 15, 184-187.

Bettleheim, B. Bringing up children. Ladies Home Journal, 1973,

Bond, G. L. and Tinker, M. A. Reading difficulties their diagnosis and correction (3rd ed.). New York: Appleton-Century-Crofts, Meredith Corporation, 1973.

Cicirelli, V. G. Sibling constellation, creativity, I.Q., and academic achievement. Child Development, 1967, 38, 481-490.

Clarke, L. Can't read, can't write, can't talk too good either. New York: Penguin Books, 1973.

Critchley, M. The dyslexic child. Springfield, Illinois: Charles C. Thomas, 1970.

Delacato, C.H. Neurological organization and reading. Springfield, Illinois: Charles C. Thomas, 1966.

Draper, N. R. and Smith, H. Applied regression analysis. New York: John Wiley and Sons, Inc., 1966.

Gillingham, A. and Stillman, B. Remedial training for children with specific disability in reading, spelling, and penmanship. Cambridge, Massachusetts: Educators Publishing Service, Inc., 1960.

Gulliland, H. A practical guide to remedial reading. Columbus, Ohio: Charles E. Merrill Publishing Company, 1974.

Hansen, H. S. The impact of the home literacy environment on reading attitude. Elementary English, 1969, 46, 17-24.

Harris, A. J. and Sipay, E. R. How to increase reading ability (6th ed.). New York: David McKay Co., Inc., 1975.

Hermann, K. Reading disability. Springfield, Illinois: Charles C. Thomas, 1959.

Hodges, A. and Balow, B. Learning disability in relation to family constellation. Journal of Educational Research, 1961, 55, 41-42.

Kerlinger, F. N. Foundations of behavioral research (2nd ed.) New York: Holt, Rinehart, and Winston, Inc., 1973.

Jencks, C. Inequality: A reassessment of the effect of family and schooling in America. New York: Basic Books, 1972.

Kaluger, G. and Kolson, C.J. Reading and learning disabilities. Columbus, Ohio: Charles E. Merrill Publishing Co., 1969.

Kennedy, E.D. Classroom approaches to remedial reading. Itasca, Illinois: F.E. Peacock Publishers, Inc., 1971.

McNeil, K. A., Kelly, F. J. and McNeil, J. T. Testing research hypotheses using multiple linear regression. Carbondale: Southern Illinois University Press, 1975.

Michael-Smith, H., Morgenstern, M. and Karp, E. Dyselxia in four siblings, Journal of Learning Disabilities, 1970, 3, 185-192.

Newman, I. A revised suggested format for the presentation of multiple regression analysis. Multiple Linear Regression Viewpoints, 1976, 6, 50-52.

Newman, I. and Benz, C.R. Multiple linear regression. U.S.A.: Multiple Regression Special Interest Group of the American Educational, 1979.

Newman, I., Deitchman, R., Burkholder, J., Sanders, R. and Ervin, L., Jr. Type IV error: Inconsistency between the statistical procedure and the research question. Multiple Linear Regression Viewpoints, 1976, 6, 1-10.

Newman, I. and Fry, J. A response to a note on multiple comparison and comment on shrinkage. Multiple Linear Regression Viewpoints, 1972, 2, 36-39.

Oberlander, M. and Jenkins, N. Birth order and academic achievement. Journal of Invididual Psychology, 1967, 23, 103-110.

Otto, W. Family position and success in reading. Reading Teacher, 1965, 19, 119-123.

Robinson, H. Why pupils fail in reading. Chicago: University Chicago Press, 1946.

Silver, L. B. Familial patterns in children with neurologically based learning disabilities. Journal of Learning Disabilities, 1971, 4, 5-15

Spache, G. D. Investigating the issues of reading disabilities. Boston: Allyn and Bacon, Inc., 1976.

Sullivan, D. D. and Humphrey, J. H. Teaching reading through motor learning, Springfield, Illinois: Charles C. Thomas, 1973.

Sutphin, F. A perceptual testing-training handbook for first grade teachers. Florida: Winterhaven Lions Club, 1967.

Symmes, J. S. and Rapoport, J.L. Unexpected reading failure. American Journal of Orthopsychiatry, 1972, 42, 82-91.

Uhl, N. and Eisenberg, T. Predicting shrinkage in the multiple correlation coefficient. Educational and Psychological Measurement, 1970, 30, 487-489.

Ward, J. H. and Jennings, E. E. Introductions to linear models. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Watson, G. Nutrition and your mind: The psychochemical response. New York: Harper and Row, 1972.

Table 1

	Models and Explanations	R^2	df	Alpha	F	p	Sign.
H_1	Model 1: $X_1 = a_0U+a_2X_2+a_3X_3+a_4X_4+a_5X_5+a_6X_6+ a_7X_7+a_8X_8+a_9X_9+a_{10}X_{10}+a_{11}X_{11}+ a_{12}X_{12}+a_{13}X_{13}+a_{14}X_{14}+a_{15}X_{15}+E$	.1912	1/79	.05	.0293	.8649	NS
	Model 2: $X_1 = a_0U+a_2X_2+a_3X_3+a_4X_4+a_5X_5+a_7X_7+ a_8X_8+a_9X_9+a_{10}X_{10}+a_{11}X_{11}+a_{12}X_{12} +a_{13}X_{13}+a_{14}X_{14}+a_{15}X_{15}+E$	.1909					
H_2	Model 1: $X_1 = a_0U+a_2X_2+a_3X_3+ \ldots a_{15}X_{15}+E$	.1912	6/79	.05	1.4424	.2091	NS
	Model 3: $X_1 = a_0U+a_2X_2+a_3X_3+a_5X_5+a_6X_6+ a_7X_7+a_8X_8+E$	.1026					
H_3	Model 1: $X_1 = a_0U+a_2X_2+a_3X_3+ \ldots a_{15}X_{15}+E$	.1912	1/79	.05	.4076	.5252	NS
	Model 4: $X_1 = a_0U+a_2X_2+a_3X_3+a_4X_4+a_5X_5+a_6X_6 +a_7X_7+ a_9X_9+a_{10}X_{10}+a_{11}X_{11}+ a_{12}X_{12}+a_{13}X_{13}+a_{14}X_{14}+a_{15}X_{15}+E$	.1870					

Note:
See page 3 for variables
S = significance - probability < .05
NS = nonsignificance - probability > .05

Table 1 (continued)

	Models and Explanations	R^2	df	Alpha	F	p	Sign.
H_4	Model 1: $X_1 = a_0U+a_2X_2+a_3X_3+ \ldots a_{15}X_{15}+E$	.1912					
	Model 5: $X_1 = a_0U+a_2X_2+a_3X_3+a_4X_4+ a_5X_5+ a_7X_7+E$	.0998	8/79	.05	1.1165	.3614	NS
H_5	Model 1: $X_1 = a_0U+a_2X_2+a_3X_3+ \ldots a_{15}X_{15}+E$	.1912					
	Model 13: $X_1 = a_0U+a_2X_2+a_3X_3+a_4X_4+a_5X_5+a_9X_9+ a_{10}X_{10}+a_{11}X_{11}+a_{12}X_{12}+a_{13}X_{13}+ a_{14}X_{14}+a_{15}X_{15}+E$	.1021	3/79	.05	2.9003	.0401	S
H_6	Model 6: $X_1 = a_0U+a_4X_4+E$	.0027					
	Model 99: $X = a_0U+E$	.0	1/90	.05	.2434	.3113	NS
H_7	Model 7: $X_1 = a_0U+a_5X_5+E$	.0240					
	Model 99: $X_1 = a_0U+E$	.0	1/90	.05	2.2142	.0701	NS

Note:
See page 3 for variables
S = significance - probability < .05
NS = nonsignificance - probability > .05

corrected
.00833
for H_6 to H_{12}

Table 1 (continued)

	Models and Explanations	R^2	df	Alpha	F	p	Sign.
H_8	Model 8: $X_1 = a_0U+a_6X_6+E$	.0740	1/90	.05	7.1904	.0044	S
	Model 99: $X_1 = a_0U+E$	.0					
H_9	Model 9: $X_1 = a_0U+a_7X_7+E$	.0826	1/90	.05	8.1004	.0028	S
	Model 99: $X_1 = a_0U+E$	.0					
H_{10}	Model 10: $X_1 = a_0U+a_8X_8+E$	.0004	1/90	.05	.0352	.4257	NS
	Model 99: $X_1 = a_0U+E$	.0					
H_{11}	Model 11: $X_4 = a_0U+a_6X_6+E$	.0237	1/90	.05	2.1890	.0712	NS
	Model 99: $X_4 = a_0U+E$	.0					
H_{12}	Model 12: $X_4 = a_0U+a_7X_7+E$	.0573	1/90	.05	5.4732	.0108	NS
	Model 99: $X_4 = a_0U+E$	.0					

Note:
See page 3 for variables
S = significance - probability < .05
NS = nonsignificance - probability > .05

corrected
.00833